DOCUMENT RESUME

ED 326 496 SP 032 723

AUTHOR Vincent, Paul
TITLE Teacher Perceptions of Maine's Master Teacher Criteria. Occasional Paper Series, No. 2.
INSTITUTION Maine Univ., Orono. Coll. of Education.
PUB DATE 87
NOTE 15p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Evaluation Criteria; *Master Teachers; *Teacher Attitudes; *Teacher Characteristics; Teacher Effectiveness; Teacher Evaluation; Teacher Qualifications; Teacher Role; *Teacher Selection

ABSTRACT

This study examines the content and nature of Maine's locally devised definitions of master teachers from three perspectives: (1) the criteria which have been proposed by 16 local committees; (2) the acceptability of such criteria to a broad cross section of Maine teachers; and (3) the comparability of what has been proposed to what appears acceptable. The 16 locally developed lists and the criteria which comprise them were examined and compared on the basis of: (1) a conceptual framework derived from categories used in the Tennessee Career Ladder Program; and (2) types of criteria used in research on teaching. The criteria content were then assessed through a survey of a random sample of Maine teachers; the analysis of the survey provided information about the content and nature of criteria which are most likely to be accepted or rejected by a broad cross section of Maine teachers. The final analysis explores the congruence between the criteria preferences of local committees and the criteria preferences of Maine teachers. Based on this analysis, implications of the study which should help the designers of future criteria lists are noted. (18 references) (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED326496

Occasional Paper Series

No. 2

Teacher Perceptions of Maine's Master Teacher Criteria

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

W. G. McIntire

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

By

Paul Vincent, Ed.D.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The Occasional Paper Series is intended to place before the educational leadership in Maine information which can assist them as they ponder solutions to complex problems confronting their communities, educational systems, or their students. Papers will be distributed periodically as topics vital to educational improvement are addressed by researchers at the University of Maine.

University of Maine

College of Education

Center for Educational Field Services

Spring 1987

SP032723

TEACHER PERCEPTIONS OF

MAINE'S MASTER TEACHER CRITERIA

Dr. Paul Vincent

Elementary Supervising Principal

M.S.A.D. #6

Paul Vincent taught in Maine schools from 1965 through 1975. He then served as elementary principal in Oakland, Maine and Superintendent of Schools in Bingham, Maine. He is presently employed as a supervising principal in M.S.A.D. #6.

Dr. Vincent received his B.S. and Ed.D. degrees from the University of Maine. His M.S. degree was earned at the University of Southern Maine. During the 1985-86 school year, he served as a research assistant for the University of Maine, College of Education. He has published articles in the *Journal of Maine Education* and has presented papers at meetings of the New England Educational Research Organization and the American Educational Research Association.

Preface

America's public education systems are in the process of coping with a dilemma created by the current educational reform movement. On one hand, many politicians, educators, and citizens agree that schemes such as merit pay and master teacher plans should be implemented to reward outstanding teachers. Twenty-six states are presently involved with plans to reward outstanding performance (Cornett and Weeks, 1985). On the other hand, professional educators have historically been unable to agree on a satisfactory definition of outstanding teaching. Teaching has been characterized as so complex and context-dependent that it may not be reducible to a definition which is widely accepted within the teaching rank.

Although the public now appears willing to provide higher salaries through merit pay, career ladders, and master teacher plans, educators may not have the ability to respond adequately. Teacher acceptance of these plans appears crucial to their success, but it may be impossible to articulate a definition of outstanding teacher which would make such plans work. It is incumbent upon the education profession to demonstrate that it can develop practical definitions which will be acceptable to those affected. The challenge will be to devise criteria of outstanding teaching which go beyond the need to satisfy the popular outcry and serve to improve the teaching profession by helping these plans succeed.

Whether the profession can respond adequately to this challenge may be demonstrated in twenty Maine school systems which are pilot testing the state's new certification law. At the time of this writing, sixteen of these systems have devised lists of criteria for identifying outstanding teachers (described as Master Teachers in the Maine law). After reviewing past efforts to define outstanding teaching, this study examines the content and nature of Maine's locally-devised definitions from three perspectives: (a) the criteria which have been proposed by local committees, (b) the acceptability of such criteria to a broad cross-section of Maine teachers, and (c) the comparability of what has been proposed to what appears acceptable.

In order to facilitate comparison and analysis of the proposed lists, a framework is developed for categorizing and classifying the content and nature of criteria. This framework is derived from the categorization scheme utilized in Tennessee's Career Ladder Program and from the type of variables commonly used in research on teaching. It serves as the basis for describing the sixteen locally developed lists and the criteria which comprise them.

The design and results of a survey which was used to assess Maine teachers' perceptions of suggested criteria are also described. The analysis of this survey provides information about the content and nature of criteria which are most likely to be accepted or rejected by a broad cross-section of Maine teachers.

The final analysis explores the congruence between the criteria preferences of local committees (with extensive teacher involvement) and the criteria preferences of Maine teachers. Based on this analysis, recommendations are made which should help the designers of future criteria lists.

The author conducted a comprehensive study of the criteria which sixteen Maine school systems have proposed for selecting master teachers. This paper summarizes the study and addresses the design of outstanding teacher selection programs in light of the study's results. It is divided into four sections: (a) a summary of the problem, methodology, and results; (b) conclusions drawn from the results; (c) limitations of the study; and (d) implications of the study.

Summary

Many of the state level educational reforms which have swept the country have included a component aimed at selecting outstanding teachers for special recognition, status, or pay. Educators have agreed that teacher acceptance of such plans is crucial to their success and that the selection criteria utilized may affect their level of acceptance. However, there has been little agreement on what constitutes a definition of outstanding teacher and little has been learned about what criteria are suitable for selecting one. In order to make recognition plans successful, we must learn more about outstanding teacher criteria and how teachers perceive them.

The study was aimed at learning more about such criteria by focusing on those proposed in sixteen Maine sites. Three research questions addressed various aspects of Maine's locally developed criteria. The first question was:

> ***What content and characteristics describe the master teacher criteria lists utilized in the Maine pilot sites?***

A detailed examination of the criteria adopted in sixteen Maine sites was conducted. Criteria and, subsequently, criteria lists were compared on: a) the basis of a conceptual framework derived from the content categories used in the Tennessee Career Ladder Program, b) the types of criteria used in research on teaching, and c) criteria inference level.

The categorization scheme used by Tennessee in its Career Ladder Program was deemed appropriate for describing criteria content. The Tennessee criteria list is familiar to many practitioners since it has been widely publicized, distributed, and discussed. It is typical of schemes which have been used to group teacher competencies and behaviors for various purposes (Vincent, 1986).

The Tennessee list is divided into six categories: planning for instruction, delivery of instruction, evaluation of student progress, classroom management, professional leadership, and basic communication skills. Most Maine lists contained criteria from all of these categories, but with varying emphasis. Items like those from Tennessee's Professional Leadership category (28%) and Classroom Management (16%) were also utilized extensively. The remaining three categories [Evaluation of Student Progress (8%), Planning for Instruction (8%), and Basic Communication Skills (7%)] were represented but accounted for a smaller portion of the lists.

Several types of criteria can be used to address the same content. The four types used in the study are based on descriptions of variables found in the literature relating to research on teaching. The process and outcome types are closely related to the process-product paradigm. They are widely used to describe the independent and dependent variables in studies of teacher effectiveness (e.g., ERS 1983; Medley 1985; Ryans 1960). The input and context types are based on the presage and context variables described by Dunkin and Biddle (1974) and Gage (1978). The following example of criteria show how the same content can be expressed in terms of the four different types.

Input criteria deal with predetermined teacher characteristics or qualities. e.g. The teacher knows how to write lesson plans.

Process criteria describe what a teacher does. e.g. The teacher writes lesson plans.

Outcome criteria deal with outcomes of a teacher's work. e.g. As a result of lesson planning, classes are well organized.

Context criteria always tie the variable to some condition such as grade level, subject area, age of students, or community expectations. e.g. The teacher writes lesson plans which are appropriate for the grade being taught.

The Maine lists were primarily composed of process criteria (89%). Input criteria were sparsely used (8%) while context specific criteria (1%) and outcome criteria (0.4%) were practically non-existent.

In addition to differences in content and type, criteria differ in specificity or inference level. High inference items are abstractly stated in broad or general terms. They cannot be rated without making a judgment. Low inference items refer to specific qualities, characteristics, behaviors, or outcomes. They are clearly observable or measurable.

Maine criteria were usually worded in high inference or abstract terms (85%). Few of them could be applied without using a great deal of rater judgment.

The comparison of the sixteen Maine lists was relatively simple, since all but two were predominantly composed of criteria of the same type (process) and inference level (high). A comparison on the basis of content, revealed four distinct list models.

The lists which fell into each model were remarkably similar to each other. The Tennessee model (4 lists) closely paralleled the original Tennessee List. The short comprehensive model (7 lists) contained fewer items than the Tennessee model but was comprised of criteria which were similar to Tennessee's. The professional leadership model (4 lists) was dominated by items from the Professional Leadership category. The classroom instruction model (1 list) contained primarily Delivery of Instruction and Planning for Instruction items.

The second research question was:

What criteria content and characteristics will a random sample of Maine Teachers accept or reject for inclusion in a definition of outstanding teacher?

This question was probed through the design and administation of a questionnaire. One hundred eighty-seven Maine teachers rated various content and criteria for acceptability and measurability. In general, these respondents indicated that most content topics were both acceptable and measurable. Nevertheless, there were differences in levels of acceptability and perceived measurability.

Content from the Delivery of Instruction, Classroom Management, and Evaluation of Student Progress categories was most acceptable to the respondents. In addition, they showed a preference for items which addressed student instructional needs, process and context type criteria, and high inference items. The least acceptable content was from the Professional Leadership, Planning for Instruction, and Basic Communication Skills Categories. Input type criteria and low inference criteria were also rated low in acceptability.

Perceived measurability was highest for low inference items, outcome type criteria, and Planning for Instruction content. Some content which was **not** included in the Tennessee list (e.g., years of experience, certification level, knowledge of subject) was also rated high in measurability. Perceived measurability was lowest for high inference items, input type criteria, and Basic Communication Skills content.

The third research question was:

To what extent will the content and characteristics which teachers approve be related to the content and characteristics of criteria utilized by the Maine pilot sites?

This question was examined in a comparison of criteria utilized in the Maine lists with those the survey respondents found to be most and least acceptable and measurable. Both congruities and incongruities were uncovered.

When utilization level was compared to acceptability levels, agreement was evident for Planning for Instruction and Basic Communication Skills. They were disfavored by both the sites and the respondents. Disagreement was evident for Professional Leadership content. It was the most frequently utilized but least acceptable category.

The sites and the respondents agreed on three criteria types. They both favored process criteria but did not favor input and outcome types. They disagreed on context criteria, which the respondents rated high.

Agreement was also evident for inference level. High inference items were extensively utilized and were more acceptable than low inference items.

Fewer congruities were evident when the perception of measurability was considered. There was only one content category and one criteria type for which local utilization, teacher acceptability, and perceived measurability levels matched. Basic Communication Skills content and input criteria received low ratings on each. These were also the only two characteristics for which there was a match between acceptability and measurability. The characteristics which were rated high in measurability were rated low in both utilization and acceptability.

Conclusions

The preceding results led to conclusions regarding: teachers' criteria preferences, the unresolved problem of measurability, and the effect of allowing criteria lists to be developed locally.

Teacher Preferences

The survey discovered some useful generalizations about teacher perceptions of criteria from different content categories and different criteria types. Three generalizations about teacher preferences are pertinent to those who wish to have their lists accepted on a broad basis.

First, the respondents agreed that most topics from the survey should be included in an outstanding teacher definition. Thirty-one of the forty-two topics were acceptable to seventy-five percent or more of the respondents. All but four items were acceptable to a majority of them. Teachers appeared to believe that an outstanding teacher would be characterized by many different qualities. No single quality or group of qualities clearly stood out as most important. This generalization was supported in the comments made by ten respondents who noted in the questionnaire margins that all categories were equally important.

Acceptability responses were also evenly distributed across content categories. All categories were rated in the acceptable range, although Professional Leadership placed last in every analysis of acceptability. This led to the inference that teachers would favor a list which placed equal emphasis on all content, except Professional Leadership and some particular items from the "Other" category (certification level and years of teaching experience).

Second, an important generalization emerged when individual criteria were examined without regard to their category. The most highly favored items related to subject matter, students, or classrooms. Respondents gave the lowest acceptability ratings to criteria that related to peripheral

issues such as (a) extra curricular activities, (b) years of teaching experience, (c) certification status, (d) physical arrangement of the classroom, and (e) use of facilities outside the classroom. In the respondents' view, outstanding teachers should be judged by what they do in their classrooms with children. They should not be judged on the basis of criteria such as those in the Professional Leadership category, which deal with activities outside the classroom in the absence of children.

This generalization was also supported in the part of the survey which assessed criteria type. In general, the respondents preferred process criteria. However, they preferred context criteria if the context referred to the students' varying needs. Clearly, criteria which related to the instructional needs of students were the most acceptable to the teachers. These criteria, therefore, should receive careful consideration in the development of outstanding teacher selection programs.

Third, when respondents compared different criteria types, outcome items were unacceptable. Items which related to student tests, assessment, and evaluation ranked last or next to last in acceptability. This presented a paradox: teachers preferred criteria which addressed student instructional needs, yet they did not favor criteria which represented the measurement of how well those needs are met. The contradiction is similar to those discussed under measurability.

Measurability

Three sections of the survey indicated that there were inconsistencies between the kinds of criteria which teachers preferred and those which they perceived as easiest to measure. Analyses of criteria content, type, and inference level pointed to several of them.

Content. The three content categories which were rated highest in acceptability were rated in the mid-range of measurability. However, many of the individual topics which were most acceptable were among those rated as most difficult to measure. Some respondents (15) emphasized this point. They commented that, in general, the qualities which were most important to being an outstanding teacher were either difficult or impossible to measure (e.g., motivation, empathy, intuitiveness).

Conversely, one of the categories (Planning for Instruction) which was rated as easiest to measure was among the least acceptable. This was also true for several particular items such as certification level and years of experience. They were rated high in measurability but near the bottom in acceptability. Thus selection of criteria is compounded by the fact that acceptable criteria were often seen as difficult to measure, and measurable criteria were often unacceptable.

Type. On one hand, several process items were rated most acceptable and easiest to measure (among four types being compared by the respondents). On the other hand, several process items which were rated high in acceptability were rated next to lowest in measurability. Conversely, four outcome items which were rated first or second in measurability were rated last or next to last in acceptability. In the section of the survey which examined criteria type, some of the most acceptable items were perceived as least measurable; and some of the least acceptable were seen as most measurable.

Inference Level. When the respondents compared high and low inference items, the most acceptable were those which they perceived as least measurable. Teachers preferred high inference items, but perceived low inference items as easier to measure. Peterson and Peterson (1984) pointed out that evaluative criteria have been criticized for being high inference items which were

not measurable. The results of this study indicated that the same criticism is likely to befall outstanding teacher criteria. The sixteen Maine lists were dominated by high inference items, which Maine teachers thought were least measurable.

In summary, it would be difficult to predict what teachers would perceive as both acceptable and measurable. Although the study found an overall positive relationship between measurability and acceptability of content and criteria type, there was one content category and many items for which acceptability and measurability were ranked at opposite extremes. For criteria type the relationship was clear; acceptability and measurability were consistently at opposite extremes. Teachers did not prefer the specific, measurable, low inference items. Teachers appeared to agree that certain kinds of criteria were easiest to measure, but they were not necessarily willing to accept them.

These results generate many questions about the feasibility of developing useful lists on the basis of teacher preferences. Criteria used in the past were criticized for being unspecific and unmeasurable, but these same criteria are the kind which teachers in the survey appear to prefer. Selection program developers who wish to use acceptable criteria will be faced with problems of measurement. Somehow, they will need to convince teachers that specific criteria and outcome criteria can be applied fairly.

Local Development

The state's decision to rely on local groups to devise master teacher criteria appears to have been based on two assumptions: first, that local list development would lead to a variety of lists and approaches; and second, that teachers would know what was acceptable to other teachers. Involving local teachers in the selection of criteria might help ensure some variability and some degree of acceptability to teachers at large.

The first assumption is not supported by the examination of such lists. All of the lists except one utilized criteria which were similar to those used in typical evaluation forms. They covered the same categories which Wise, Darling-Hammond, McLaughlin, and Berstein (1985) and Lewis (1982) found in their reviews of teacher evaluation procedures. The local lists were similar in their uneven emphasis of the categories utilized. Professional Leadership items dominated most of the lists. In addition, there was little variation in the way criteria were written. Eighty percent of all items were of the process type, and eighty-one percent were high inference criteria.

Relying on local groups to devise master teacher criteria lists did not lead to a great deal of variability from site to site. On the basis of this study, fifteen of the sixteen Maine lists appeared to be more similar than they were different. This was particularly evident in the common reliance on high inference, process type, Professional Leadership criteria.

The second assumption was supported by many authors, but it was not entirely supported by this study. Weeks (1985), Schlechty, Joslin, Leak, and Hanes (1985), Riddle (1925), and Burgess (1984), for example, all encouraged teacher involvement in criteria development. This study did not refute claims that teacher involvement in a local process would increase the shared understanding which Darling-Hammond, Wise, and Pease (1983) felt was necessary. Teacher participation in criteria selection might increase acceptability for those who were directly involved, but it might not have an effect on teachers who were not.

The Maine lists were developed with extensive teacher involvement. Teachers participated in the development of all sixteen lists analyzed in this study. For most of the lists, teachers comprised at least a majority of the criteria selection group. However, the resulting criteria lists did not emphasize the content which was most acceptable to the teachers surveyed. Although most content was acceptable to the respondents, the content which they favored least received the greatest emphasis in the local lists.

Limitations

The findings and conclusions which are discussed above provide considerable information for criteria list developers. Before examining the implications of these results, it is necessary to inform the reader of several limitations. First, the criteria lists were treated as though they were an exhaustive representation of the site's master teacher concept. However, use of selection procedures which included (a) interviews, (b) recommendations from other professionals, (c) personal portfolios, and (d) classroom observations indicated that unlisted factors may have been an important part of the local definitions. These factors were not assessed in the survey. The reader must be cautioned about assuming that the lists represent all aspects of the teacher selection process.

Second, while the sample appeared to represent the demographic characteristics of the population of Maine teachers, the response rate was 37%. Although one can only speculate about the non-respondents, it is clear that the respondent group does not include teachers who were not motivated by the topic and the survey. Whether this depressed or inflated acceptability scores remains open to conjecture. It is reasonable to assume that responses would have varied more widely if the opinions of unwilling respondents had been assessed.

Third, the reliability of those parts of the survey which examined criteria items may have been affected by the wording of the items. When a criteria characteristic was being assessed, every attempt was made to keep the other two characteristics constant; but, this may not have been completely successful.

Finally, the survey results were more homogeneous than expected. This led to the important findings that teachers were willing to accept a broad variety of content and viewed most content as measurable. On the other hand, this homogeneity made it difficult to distinguish between levels of acceptability and measurability.

Implications

The major conclusions discussed above imply policy or procedural considerations for practitioners. They provide a basis for further work in three areas: (a) developing an acceptable outstanding teacher criteria list; (b) addressing the problems associated with measurability; and (c) involving local educators in list development.

List Acceptability

It appears from the survey results that teachers would favor a broad, comprehensive list which covers all aspects of teaching. Nearly all of the content presented in the questionnaire was judged acceptable. Teachers appeared to feel that an outstanding teacher was outstanding in all areas surveyed. If more areas had been surveyed, those too might have been judged important.

In their comments, some respondents suggested that the questionnaire did not, and could not, cover all aspects of outstanding teaching. Sample comments included:

"There are too many variables to be listed."

"I don't sense a recognition here of the complexity of the issue."

"This questionnaire is far removed from the day to day occurrences and experiences of teachers in the classroom."

For some respondents, the whole list seemed off target. They saw outstanding teaching as more intangible than suggested by the questionnaire items. For example:

"Teachers have been endowed with special caring and talents from above."

"How do you critique intuitiveness, instincts, and old fashioned gut feeling which we subconsciously employ in the classroom?"

"... a lot of **magic** goes unrecognized."

In 1925, Anna Riddle indicated that teachers thought it was impossible to measure the "whole value" of a teacher's work. The survey results imply that many teachers still think that no list can represent **all** of the characteristics of an outstanding teacher. Teachers with this view are bound to be displeased with any list proposed.

Furthermore, the homogeneity of the respondents' acceptability ratings indicates that it would be a mistake to use any list which emphasized one topic over others. It is clear that teachers in the sample did not want to be rated on Professional Leadership criteria or other items which related to what they did outside of their classrooms or in the absence of students. However, except for a few unpopular items, all content appears to be almost equally important to teachers.

In summary, the survey results indicate that in order to be accepted by teachers, an outstanding teacher criteria list would need to (a) cover a broad range of topics, (b) place equal emphasis on **most** topics, (c) avoid items like those from the Professional Leadership category, and (d) whenever possible include items which relate directly to students.

Measurability

The respondents ranked few topics high in both acceptability and measurability. Teachers often saw the most important items as difficult to measure. If list developers selected outstanding teacher criteria on the basis of what teachers would accept, they would end up with many unmeasurable items. Conversely, if they chose measurable criteria, they would have many unacceptable items.

This paradox will not be easily resolved. List developers will need to work closely with teachers if they wish to arrive at a list which is both acceptable and measurable. This study did point out a few items which were acceptable and measurable to most teachers. List developers could begin with these particular items in order to earn the trust of participating teachers.

The application of criteria has been a major problem in past recognition plans that failed. This study may help implementors select acceptable criteria, but this would only be the first step. They would also need to develop evaluation procedures which were more acceptable than those used in the past. In order to make these procedures work, it appears that low inference and outcome criteria should be used. Low inference criteria have been recognized by evaluation critics and by the teachers in this study as the most measurable kind. Unless they are utilized, application of criteria will be subject to rater bias and cries of unfairness will undoubtedly be

heard. Outcome criteria are needed in order to satisfy the calls for accountability. Education critics and reformers, as well, often focus on student outcome criteria such as achievement test scores.

The problem with outcome and low inference criteria, however, is that they are the least favored by teachers. Outcome items which would identify outstanding teachers on the basis of test results and demonstrated student achievement were among the least acceptable. Likewise, low inference items were consistently rated low in acceptability. Cynics might argue that teachers dislike outcome and low inference criteria because they do not want to be held accountable for their work. The literature suggests otherwise; teachers probably feared these criteria because they did not feel that they would be interpreted fairly. If these kinds of items are to be successfully used, implementors must convince teachers that criteria will be applied impartially and objectively.

Another reason to utilize low inference items is the need to make selection procedures defensible. If the selection plan involves monetary rewards or promotions, legal challenges and administrative appeals can be expected from teachers who feel they are unjustly denied recognition. To defend their selection procedures, those responsible will have to show that their criteria are objective and measurable. Unless the criteria are low inference, this will be difficult to do.

Given the many potential conflicts between acceptability and measurability which this study uncovered, the profession may not be prepared to deal with the recognition of certain teachers as more outstanding than others. Before recognition plans can be acceptable to teachers, it appears that the paradox of acceptability and measurability needs to be resolved. This study indicates that teachers would accept a definition based on what they do in their classrooms with children. However, they are not yet convinced that these processes can be measured objectively.

Local List Development

One respondent commented that a "just and fair" list could not be devised because "situations and conditions are so variable across the state." Anticipation of such criticism was probably the main impetus for allowing local list development during the pilot period. Allowing local pilot sites to develop their own criteria lists was expected to result in lists which varied according to local needs. However, the sixteen lists were not very different from each other or from lists produced in other places for teacher evaluation purposes.

In several sites, this similarity was the obvious result of heavy reliance on the Tennessee model. In other sites, it may have reflected the difficulty of defining outstanding teaching without relying on some already familiar base. The language of typical teacher evaluation forms or of effective teaching literature may have been the only common base for developing master teacher criteria lists.

Another factor contributing to this uniformity may have been the Master Teacher Certification pilot process. Each pilot site was charged with the establishment of teacher support teams which would be partially staffed by master teachers. The master teachers on these teams would act as mentors for other teachers who were being supported for recertification purposes. The heavy reliance on Professional Leadership and Delivery of Instruction criteria may have occurred in anticipation of the mentoring role which master teachers would be expected to play. These kinds of criteria may have been chosen because they reflected the skills needed to help other teachers in their classrooms.

This presents an interesting dilemma. If master teachers are selected to fill a new role in the schools, the criteria used to select them should be related to that role. However, these may not be the same criteria which would identify the persons who are held in highest regard by other teachers. The most outstanding teacher may not necessarily be the one who possesses the skills needed to work with other teachers as a mentor, evaluator, staff developer, or curriculum leader. Criteria for recognizing outstanding teachers may need to be different from criteria for selecting teachers to fill new roles such as those proposed under career ladder plans.

Extensive teacher participation in the selection of Maine's criteria was expected to increase their acceptability to other teachers. However, it did not result in the selection of the most acceptable list content identified by the survey. As pointed out previously, the sixteen similar lists emphasized professional leadership criteria which were least acceptable to a broad cross-section of Maine teachers.

The ultimate contribution of local list development may not have been originality or variability. It may have simply been the involvement of teachers. It may have been a way of giving teachers a feeling of control over their own profession. Time and effort could have been saved if a statewide list of acceptable criteria had been developed and disseminated, but this approach would have reduced opportunities for participation at the local level.

Summary of Implications

It will be difficult to devise a list of criteria which completely defines an outstanding teacher. Teachers would expect such a definition to be extensive and possibly exhaustive. They see many criteria as equally important.

In addition, many of the qualities required of an outstanding teacher are very low in perceived measurability. In particular, when an important and acceptable factor is reduced to a low inference measurable criterion, it is likely to become unacceptable.

Local list development will not necessarily lead to variability among lists, nor will it necessarily lead to criteria which are acceptable to a broad cross-section of teachers. It may, however, help garner teacher support by giving them a role in their own recognition program.

Fortunately, some of the criteria content which were needed and adopted by Maine's pilot sites appeared to be both acceptable and measurable. These topics may provide a good starting point for future efforts. Nevertheless, this study indicated that problems of measurability will need continuing attention. Extensive collaboration between teachers and program developers will be needed if criteria lists are to improve and if application procedures are to be perceived as fair.

Bibliography

Burgess, C.E. (1984, September). Close-up on merit pay. *Learning,* 55-60. Career Ladder Clearing House. (1984, December). News from the states. Atlanta: Southern Regional Education Board.

Cornett, L., & Weeks, K. (1985, July). Career ladder plans: Trends and emerging issues 1985. Atlanta: Career Ladder Clearing house, Southern Regional Education Board.

Darling-Hammond, L., Wise, A.E., & Pease, S.R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, 53*(3), 285-328.

Dunkin, M., & Biddle, B.J. (1974). *The study of teaching.* New York: Holt, Rinehart and Winston.

Educational Research Service. (1983). *Merit pay for teachers: Status and description.* Arlington, VA.

Gage, N.L. (1978). *The scientific basis of the art of teaching.* New York: Teachers College Press.

Lewis, A.C. (1982). *Evaluating educational personnel.* Arlington, VA: American Association of School Administrators.

Medley, D.M. (1985). Evaluating teaching: Criteria. In T. Husen, (Ed.), *The international encyclopedia of education.* New York: Pergamon Press.

Peterson, D., & Peterson, K. (1984, February). A research based approach to teacher evaluation. *NASSP Bulletin, 68*(469), 39-46.

OP9471:2/galley 10

Riddle, A. (1925). Report of the committee of one hundred on classroom teacher's problems: teacher rating. *Addresses and Proceedings of the National Education Association, 63,* 200-215.

Ryans, D.G. (1960). *Characteristics of teachers: Their description, comparison, and appraisal.* Washington, DC: American Council on Education.

Schlechty, P.C., Joslin, A., Leak, S., & Hanes, R. (1985, December). The Charlotte Macklenburg teacher career development program. *Educational Leadership, 42*(4), 4-8.

Tennessee Department of Education. (1984). *Teacher orientation manual.* Nashville, TN.

Vincent, P.R., & McIntire, W.G. (1986, May). Emerging master teacher criteria. Paper presented at Annual Meeting of New England Educational Research Organization, Rockland, Maine.

Vincent, P.R. (1986, August). Teacher perceptions of criteria for identifying master teachers in Maine. Doctoral dissertation, University of Maine.

Weeks, K. (1985, March). Planning career ladders: Lessons from the states. Nashville, TN: Career Ladder Clearing House.

Wise, A., Darling-Hammond, L., McLaughlin, M.W., & Bernstein, H. (Rand Corporation). (1985, September). Teacher evaluation: A study of effective practices. *Elementary School Journal, 86*(1), 61-121.